Supplementary Material

Structural basis for conformational switching and GTP loading of the large G protein atlastin

Laura J. Byrnes[a], Avtar Singh[b], Kylan Szeto[c], Nicole M. Benvin[d], John P. O'Donnell[a], Warren R. Zipfel[b], and Holger Sondermann[a, 1]

Supplementary Materials and Methods:

Protein expression and purification

The cytoplasmic domain (residues 1-446), G domain (1-339), middle domain (340-446), truncated middle domain (1-366), and C-terminal ECFP/EYFP fusions (atlastin1 1-446 followed by a short linker containing amino acid sequence GSTSTG followed by either ECFP or EYFP) of human atlastin-1 were amplified by standard PCR and cloned into a modified pET28a expression plasmid (Novagen) yielding N-terminally hexahistidine-tagged SUMO fusion proteins. The hexahistidine-tagged SUMO-moiety was cleavable using the protease Ulp-1 from *S. cerevisiae*. The cytoplasmic domain (1-446) used for crystallization of wild-type atlastin was cloned into pET21, yielding C-terminally hexahistidine-tagged protein, which was not cleaved during purification.

Proteins were overexpressed in *E. coli* BL21 (DE3) (Novagen) or T7 Crystal Express (NEB) cells, respectively. For the expression of native proteins, cells were grown

in Terrific Broth (TB) media supplemented with 50 μg/ml kanamycin (for expression from pET28) or 100 μg/ml ampicillin (for expression from pET21) at 37°C. At an optical density corresponding to an absorbance of 0.8-1.0 at 600 nm (OD_{600}), the temperature was reduced to 18°C, and protein production was induced with 0.5 mM IPTG. After 16 hours, cells were harvested by centrifugation, resuspended in NiNTA buffer A (25 mM Tris-Cl, pH 8.5, 500 mM NaCl and 20 mM imidazole), and flash-frozen in liquid nitrogen.

After cell lysis by sonication and removal of cell debris by centrifugation, clear lysates were loaded onto NiNTA Superflow (Qiagen) equilibrated in NiNTA buffer A. The resin was washed with 20 column volumes of NiNTA buffer A, and proteins were eluted three times with 2 column volumes of NiNTA buffer A supplemented with 500 mM imidazole. Proteins were buffer exchanged into desalting buffer (25 mM Tris-HCl, pH 7.5, 400 mM NaCl, 5mM β-mercaptoethanol), with proteins expressed in pET21 constructs immediately subjected to size exclusion chromatography. In the case of proteins expressed with a SUMO moiety, affinity tags were removed by incubation with the yeast protease Ulp-1 at 4°C overnight. Cleaved proteins were collected in the flow-through during NiNTA affinity chromatography, and were subjected to size exclusion chromatography on a Superdex 200 column (GE Healthcare) equilibrated in gel filtration buffer (25 mM Tris-HCl, pH 7.5, 100 mM NaCl). Proteins were concentrated on a Centricon ultrafiltration device (10 kDa cutoff; Millipore) to a final concentration of approximately 0.5-1 mM. Protein aliquots were flash frozen in liquid nitrogen and stored at -80°C.

Site-directed mutagenesis was carried out using Quikchange (Agilent) following the manufacturer's instructions, followed by validation through DNA sequencing.

Supplementary Figure Legends

Supplementary Figure 1: Crystal structure of atlastin-1^{1-446}-N^{440}T form 3 crystal structure bound to GppNHp or GDP•AlF_4^-.
A. Overview of asymmetric unit. Atlastin-1^{1-446}-N^{440}T crystallizes as a tetramer in the asymmetric unit. An interaction of the C-terminal tail of one protomer with the G domain of an adjacent molecule is shown in the inset.
B. Detailed view of interactions of the C-terminal tail with an adjacent G domain. Residues 440-446 (grey) from molecule B and residues from the G domain of molecule C (yellow) that interact with them are shown as sticks and labeled. A F_o-F_c omit map for the tail motif is contoured at 3.5 sigma.
C. Superposition of crystal form 1 and 3. The respective G domains of crystal form 1 and 3 were superimposed to assess whether tail packing interactions coincide with conformational changes in the G domain. No significant changes were observed.
D. Nucleotide binding pocket of form 3 crystal structures. The protein is shown in cartoon presentation, GDP•AlF_4^- (upper panel) or GppNHp (lower panel) are shown as sticks, and Mg^{2+} and waters are shown as green and red spheres, respectively. Nucleotide F_o-F_c omit map is contoured at 4.0 sigma.
E. Catalytic residues in the nucleotide binding pocket of form 3 crystal structures. The protein is shown in cartoon presentation except for residues that interact with the phosphates of the nucleotide, which are shown as sticks. GDP•AlF_4^- (upper panel) or GppNHp (lower) are shown as sticks, and Mg^{2+} is shown as spheres.

Supplementary Figure 2: Crystal packing interactions.
A. Asymmetric unit and crystal packing of atlastin1^{1-446}-N^{440}T bound to GppNHp or GDP•AlF_4^-. The asymmetric unit contains 4 protomers, forming an anti-parallel dimer of dimers.
B. Asymmetric unit and crystal packing of atlastin1^{1-446}-C-His_6 bound to GppNHp or GDP•AlF_4^-. The asymmetric unit contains 2 protomers that form a very similar dimer as observed in atlastin1^{1-446}-N^{440}T structures bound to the same nucleotides.

Supplementary Figure 3: Atlastin-1^{1-446}-N^{440}T oligomerization in solution.
A. SEC-MALS data for wild-type Atlastin-1^{1-446}. The signal from the 90°-light scattering detector and refractive index detector are shown as colored, solid lines (apo, red; GppNHp-bound, green; GDP-bound, purple; GDP•AlF_x-bound, orange) and black, dashed lines respectively (left Y axis). Average molecular weight calculations across the protein peak are shown as black circles (right Y axis). The theoretical molecular weight (based on primary sequence) for the monomer and dimer are shown as horizontal dashed lines. Proteins (30-40 μM) were incubated with nucleotides (2 mM) at least 30 min prior to SEC-MALS analysis.
B. SEC-MALS data for the N^{440}T mutant introduced into atlastin-1^{1-446}.

Supplementary Figure 4: FRET in an alternative GTP analog and middle domain interaction characterization.

A. Middle and G domain time-resolved FRET following mixing with the GTP analog GTPγS. A mixture (1uM final) of atlastin-1^{1-446}-ECFP/EYFP fusion proteins (left panel) or Alexa 488 and 647-labeled atlastin-1^{1-446} (right panel) was prepared in the absence of nucleotide, and mixed 1:1 with nucleotide-containing buffer (GTPγS; 2 mM) using a stopped-flow apparatus. Experimental details are as stated in Figures 5A and 7. Time scale of FRET change was monitored over 10 min, and proceeds with similar kinetics to the GTP analog GppNHp (see Figures 5A and 7 for comparison).

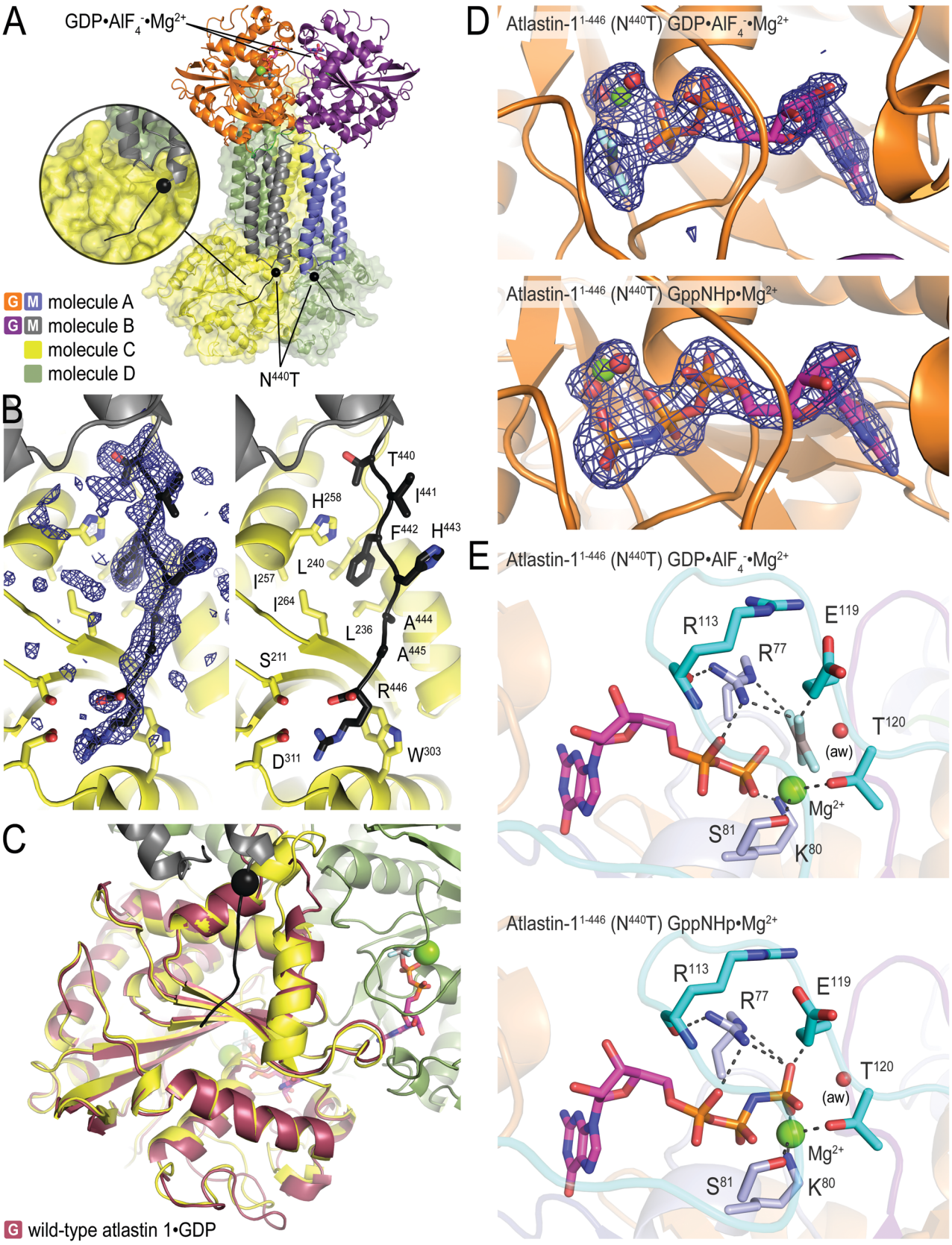
B. Middle domain titration increases G domain activity. With a G domain concentration of 2μM, middle domain concentrations (either wild-type [squares] or point mutant $M^{347}E$ [triangles]) were varied from 0.5μM to 500μM and the resulting change in activity compared to G domain alone was plotted versus the concentration of middle domain added on a $\log_{10}$ scale. The data were fit to a one-site saturated binding equation with a variable hill coefficient (fit shown in grey for wild-type; no fit could be determined for $M^{347}E$ data). The fit resulted in an apparent K_d of 62.2±6.6 μM and a hill coefficient of 0.931±0.048. Experiments were conducted in triplicate. We report the means with the error bars representing SEM.

C. Middle domain point mutation $M^{347}E$ reduces the GTPase activity of the N-terminal cytoplasmic domains of atlastin-1 (residues 1-446). The $M^{347}E$ mutation was introduced into the soluble atlastin-1^{1-446} construct, and its GTPase activity was determined by measuring the production of inorganic phosphate over time upon GTP hydrolysis at various protein concentrations.

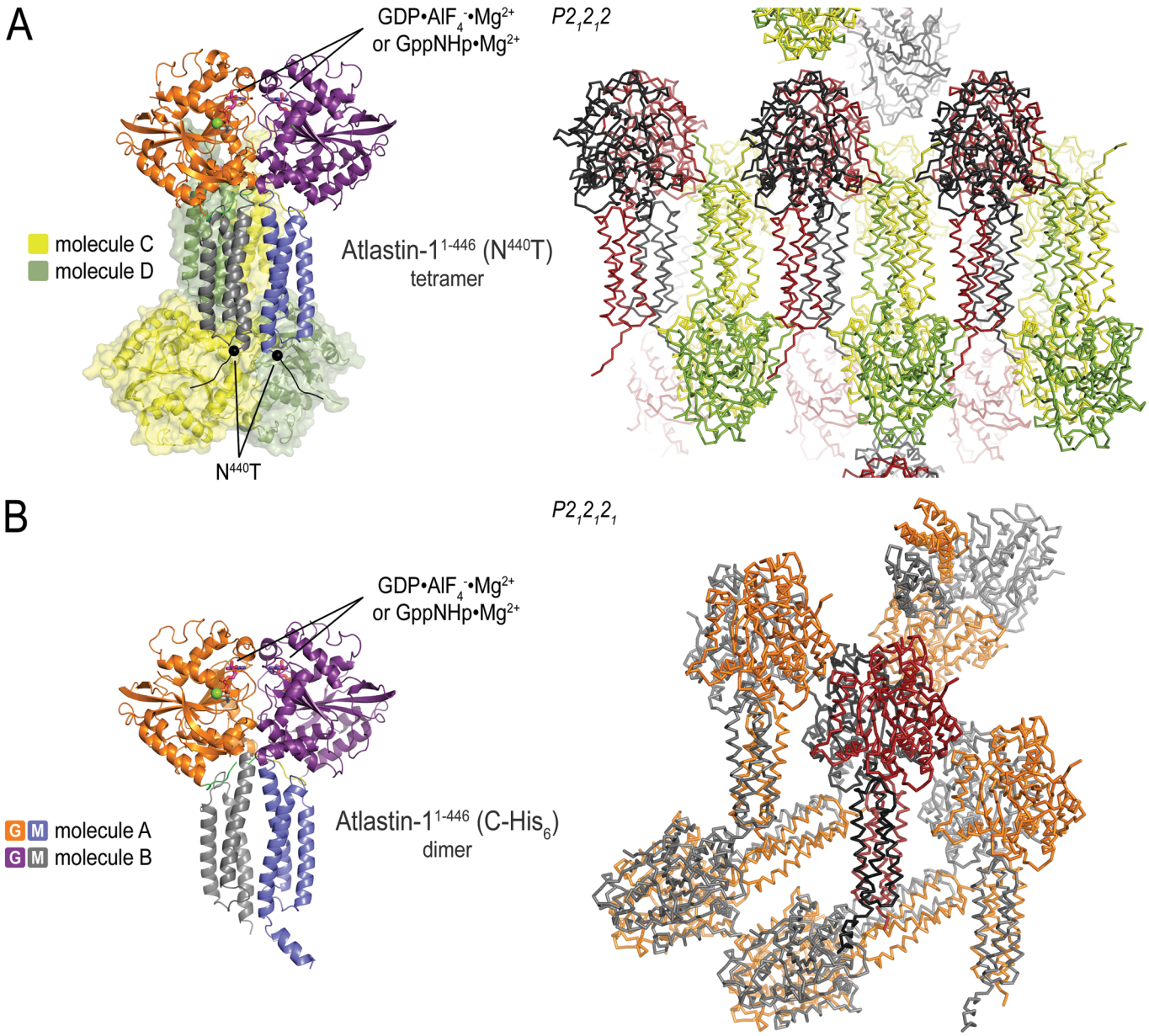
Supplementary Appendix: Modeling of the time-resolved FRET data.

Supplementary Table 1: Data collection and refinement statistics.

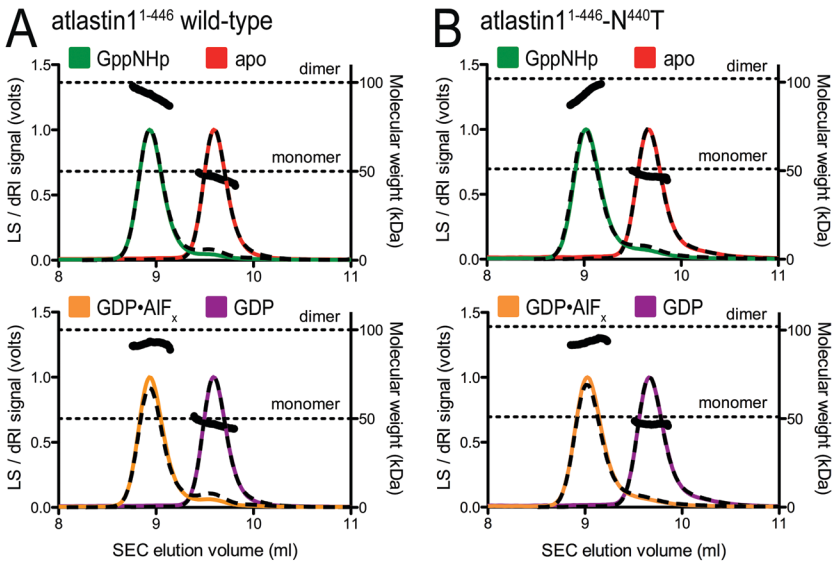
Supplementary Table 2: Comparison of G and middle domains in atlastin-1 crystal structures.



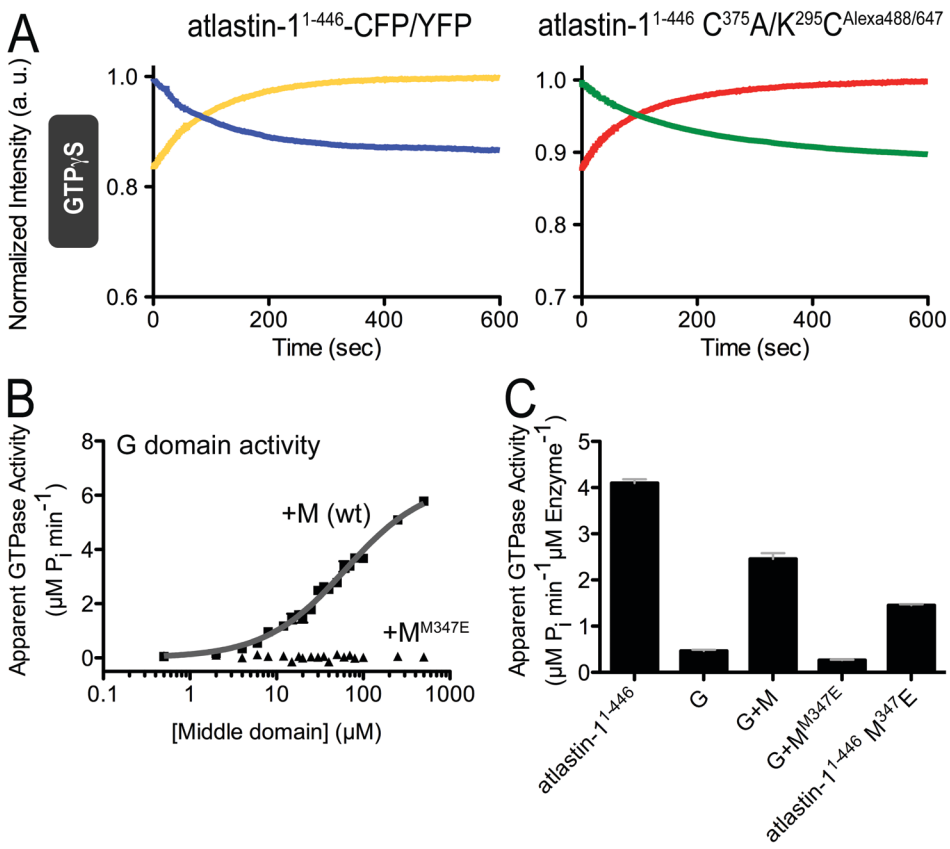
Supplementary Figure 1



Supplementary Figure 2



Supplementary Figure 3



Supplementary Figure 4

Supplementary Table 1: X-ray data collection and refinement statistics.

	Human atlastin-1 1-446, N440T GDP•AlF_4^-	Human atlastin-1 1-446, N440T GppNHp	Human atlastin-1 wild-type, C-His_6 GDP•AlF_4^-	Human atlastin-1 wild-type, C-His_6 GppNHp
X-ray source	CHESS, A1	CHESS, A1	CHESS, A1	CHESS, A1
Wavelength (Å)	0.9771	0.9771	0.9771	0.9771
Space group	$P2_12_12$	$P2_12_12$	$P2_12_12_1$	$P2_12_12_1$
Unit cell				
a, b, c (Å)	129.0, 267.1, 62.1	132.0, 268.1, 62.4	49.6, 116.4, 185.4	49.7, 115.8, 181.1
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Resolution (Å)[a]	50-2.3 (2.38-2.30)	50-2.6 (2.69-2.60)	50-2.1 (2.18-2.09)	50-2.2 (2.28-2.20)
No. of reflections				
Total	743,898 (45,513)	540,312 (44,595)	512,721 (43,745)	334,053 (9,968)
Unique	92,313 (7,714)	69,639 (6,656)	63,782 (6,730)	47,929 (2,167)
Completeness (%)	95.3 (81.1)	99.7 (97.4)	98.9 (96.8)	90.3 (41.7)
Redundancy	14.6 (3.0)	7.8 (6.7)	8.0 (6.5)	7.0 (4.6)
I/σ(I)	35.6 (10.5)	12.3 (3.1)	23.3 (4.1)	16.1 (1.9)
R_{meas} (%)	12.0 (50.4)	16.0 (61.1)	8.1 (47.4)	10.1 (59.3)
Refinement				
R_{work} / R_{free} (%)	20.6 / 25.6	20.1 / 23.6	19.9 / 24.9	17.3 / 21.8
rms deviations				
Bond length (Å)	0.008	0.014	0.008	0.006
Bond angles (°)	1.19	1.68	1.23	0.98
No. of atoms	14,506	14,015	7459	7192
Protein	13,486	13,480	6730	6762
Ligands	212	132	68	66
Water	704	403	661	364
Ave. B-factors (Å^2)				
Protein	27.8	28.2	30.7	41.5
Water	32.8	24.5	34.0	40.6
Ramachandram (%)				
Favored	99	98	98	98
Outliers	0	0	0	0.1

(a) Values in brackets are for the highest resolution bin.

Supplementary Table 2. Comparison of G and middle domains in atlastin-1 crystal structures.

Protomer RMSD:

	C-His GppNHp	C-His GDP•AlF_4^-	N440T GppNHp	N440T GDP•AlF_4^-
C-His GppNHp	X	0.209	0.268	0.271
C-His GDP•AlF_4^-	0.209	X	0.276	0.265
N440T GppNHp	0.268	0.276	X	0.157
N440T GDP•AlF_4^-	0.271	0.265	0.157	X

Dimer RMSD:

	C-His GppNHp	C-His GDP•AlF_4^-	N440T GppNHp	N440T GDP•AlF_4^-
C-His GppNHp	X	0.220	0.301	0.320
C-His GDP•AlF_4^-	0.220	X	0.335	0.310
N440T GppNHp	0.301	0.335	X	0.163
N440T GDP•AlF_4^-	0.320	0.310	0.163	X

Supplementary Appendix: **Modeling of the time-resolved FRET data.**

Dimerization Kinetics

For time-dependent FRET simulations, we propose a simple three-state model based on the available structural information: a monomeric non-FRET state M, an initial dimerization state D_1, and a possible "relaxed" dimer state D_2. We assume that the dynamics of nucleotide binding is sufficiently fast as to not significantly contribute to apparent dimerization and FRET behavior and is therefore ignored. Dimerization is modeled by the 2nd order kinetic equation:

$$\frac{d[D]}{dt} = k_{on}[M]^2 - k_{off}[D]$$

Where [D] is the dimer concentration, [M] is concentration of free monomers, k_{on} is the rate constant for dimerization, and k_{off} is the rate constant for dissociation. The two populations are related by the mass conservation equation:

$$[M_0] = [M] + 2[D]$$

Where $[M_0]$ is the initial monomer concentration at time equals 0 and the initial dimer concentration ($[D_0]$) equals 0. The monomer concentration as a function of time ([M(t)]) can be modeled numerically by discretizing the kinetic equation above:

$$\Delta[D(\Delta t_n)] = \left(k_{on}[M(\Delta t_n)]^2 - k_{off}[D(\Delta t_n)]\right)\Delta t$$

$$[M(\Delta t_{n+1})] = [M(\Delta t_n)] - 2\Delta[D(\Delta t_n)]$$

Where Δt is the discretized time interval, and the subscript n indicates the n^{th} interval such that $t = n\Delta t$. The dimer concentration [D] is modeled similarly.

However, with the incorporation of a second possible "relaxed" dimer state, the populations are determined by:

$$[D_1(\Delta t_{n+1})] = [D_1(\Delta t_n)]\left(1 - k_{off}\Delta t - k_{12}\Delta t\right) + [M(\Delta t_n)]^2 k_{on}\Delta t$$

$$[D_2(\Delta t_{n+1})] = [D_2(\Delta t_n)]\left(1 - k_{off}\Delta t\right) + [D_1(\Delta t_n)]k_{12}\Delta t$$

Where D_1 is the initial dimerization state and hence receives all new dimerization events at each time interval. D_2 is the relaxed dimer state and is acquired through D_1 at the relaxation rate k_{12}. A second, backward rate constant k_{21} could have also been incorporated, but was found to be an unnecessary parameter for simulating the observed FRET signal.

The above equations are sufficient for instantaneous transitions between D_1 and D_2. However, for a finite relaxation time, an additional set of transition states $\mathbf{D_{12}}$ is acquired with a total transition time of $\tau_{12} = m\Delta t$. Here, m is an integer representing the number of concentration elements of the vector $[\mathbf{D_{12}}]$. The final time-dependent equations can be written:

$$[D_1(\Delta t_{n+1})] = [D_1(\Delta t_n)]\left(1 - k_{off}\Delta t - k_{12}\Delta t\right) + [M(\Delta t_n)]^2 k_{on}\Delta t$$

$$[D_{12}(i, \Delta t_{n+1})] = \begin{cases} [D_1(\Delta t_n)]k_{12}\Delta t, & i = 1 \\ [D_{12}(i-1, \Delta t_n)](1 - k_{off}\Delta t), & 2 \le i \le m \end{cases}$$

$$[D_2(\Delta t_{n+1})] = [D_2(\Delta t_n)]\left(1 - k_{off}\Delta t\right) + [D_{12}(m, \Delta t_n)]k_{12}\Delta t$$

Where i indicates the i^{th} element of the set of states $\mathbf{D_{12}}$ and the total time a subpopulation of dimers has spent in transition is $\tau = i\Delta t$.

FRET Calculations

FRET can be determined by applying the Förster equation:

$$E = \frac{1}{1 + \left(\frac{r}{R_0}\right)^6}$$

Where *E* is the FRET efficiency, r is the distance between the donor and the acceptor, and R_0 is the Förster distance at which $E = \frac{1}{2}$. If the fraction of donor molecules is *F*, and acceptor molecules is (*1-F*), then the distribution of dimer pairs is given by the binomial distribution:

$$\binom{2}{k} F^k (1-F)^{2-k} \quad for\ k = 0,1,2$$

For any sub-population of dimers, the donor fluorescence (S_{DONOR}) and the FRET signal (S_{FRET}) are proportional to the probability of getting a FRET pair:

$$P_{AD} = 2F(1-F)$$

$$S_{FRET} \propto P_{AD} E(r)[D(r)]$$

$$S_{DONOR} \propto P_{AD}\big(1 - E(r)\big)[D(r)]$$

By assigning a value for *r* for each state D_1, D_2, and a set of r-values for the vector of states $\mathbf{D_{12}}$, we can determine the FRET efficiency at each state (E_1, E_2, $\mathbf{E_{12}}$). The total signal is given by:

$$S_{FRET}(\Delta t_n) = 2F(1-F)\left\{E_1[D_1(\Delta t_n)] + E_2[D_2(\Delta t_n)] + \sum_{i=1}^{m} E_{12}(i)[D_{12}(i, \Delta t_n)]\right\}$$

$$S_{DONOR}(\Delta t_n)$$

$$= F[M(\Delta t_n)] + 2F^2[D(\Delta t_n)]$$

$$+ 2F(1-F)\left\{(1-E_1)[D_1(\Delta t_n)] + (1-E_2)[D_2(\Delta t_n)] + \sum_{i=1}^{m}\big(1 - E_{12}(i)\big)[D_{12}(i, \Delta t_n)]\right\}$$

Simulation Parameters

For the GTP-binding dimer simulations, the Förster distance R_0 we used was the calculated value of 47 Å. The value of *r* for D_1 and D_2 were 30 Å and 50 Å as determined

from structural data. We estimate that the equilibrium dissociation constant for dimerization to be better than 1μM, therefore we assumed a dissociation constant $K_D = \frac{k_{off}}{k_{on}}$ of 500 nM. Assuming a k_{on} of about 0.32 $\mu M^{-1}s^{-1}$ k_{off} was calculated to be 0.16 s^{-1}. The relaxation rate k_{12} was taken to be 0.50 s^{-1} with a transition time τ_{12} of 0.10 s. The time interval Δt was 0.01 s, which matches the time resolution of the relevant instrumentation. Finally the transition states $\mathbf{D_{12}}$ were calculated by assuming a uniform transition speed:

$$\frac{\Delta r}{\Delta \tau} = \frac{D_2 - D_1}{\tau_{12}}$$

$$D_{12}(r_i) = \frac{D_2 - D_1}{\tau_{12}} i\Delta t \quad for\ i = 1 \dots m$$

For the GppNHp-binding dimer simulations, only the kinetic parameters are assumed to change. Since this system cannot undergo hydrolysis, it is assumed that the relaxation rate constant k_{12} is zero and hence does not undergo any transition (i.e. $\mathbf{D_{12}}$ and D_2 are always zero). We also observed a significantly longer time scale for equilibration of over an order of magnitude. Assuming the equilibrium dissociation constant for dimerization is unchanged, we chose to use a lower k_{on} of 0.006 $\mu M^{-1}s^{-1}$ and thus a calculated k_{off} of 0.003 s^{-1}.